Experimental Design

Experiments are used to study causal relationships (rather then "just" correlations as you can infer from observational studies)

So what are some key characteristics of a good experimental design?

Experimental Design

- Context:
  - Experiment is embedded in theory
  - Provides results that are easy to interpret
  - Permits conclusions that have wide validity (large sphere of inference)

- Design
  - Provides an unbiased estimate of effects
  - Provides information for every outcome (good effect size)

- Analysis
  - Is clearly linked to a statistical treatment

Context

- What is your research question?
- How does it fit with theory?
- Consider your variables and how they are related
- What are you going to measure?
- Write a specific, testable hypothesis
  - Brief, clear, precise, and definitive
  - States an expected relationship between two or more variables
  - Testable

Context - Why is it important that it fits within a larger theory?

Statistical errors

	There is no actual difference	There is an actual difference
Accept null hypothesis of no difference	Correct conclusion	Type II error Accept incorrect H_0 False negative
Reject null hypothesis	Type I error Reject correct H_0 False positive	Correct conclusion

Context - Why is it important that it fits within a larger theory?

What is the chance of a false positive with $\alpha=0.05$ (or 5%)?

The expectation is that 1/20 experiments where there is NO underlying difference between the treatment/populations will produce a false positive.

However, if we have a good theoretical expectation, most of our treatments are expected to produce different results. So overall the percentage of false positives is much lower than 1/20!

But if you randomly test variables, you will get 1/20 false positives.

Experimental Design

Key elements:

- **Control/comparison group**
  - A control is designed to isolate the variable of interest. It should assure that only the tested variable changes between the experimental and comparison group
  - Negative/positive control
- Replication
- Randomisation

- Experimenter bias

Experimental Design

Key elements:

- Control/comparison group
- **Replication** → is the repetition of an experimental condition
  - Reduce and estimate variation
  - Experimental unit / **independence**
- Randomisation

- Experimenter bias

Experimental Design

Key elements:

- Control/comparison group
- Replication
- **Randomisation**
  - eliminates bias
  - balances the groups with respect to known **and** unknown confounding variables
  - blocking

- Experimenter bias

Experimental Design – blocking

assign 10 animals to two treatments

Treatment A	Treatment B		Treatment A	Treatment B	Treatment A	Treatment B
1	2		2	1	1	2
3	5		4	3	4	3
4	8		5	7	6	5
6			6	8	7	8
7			9	10	10	9
9						
10						

Random

(10 random events)

Random

With a constraint of equal numbers
(5 random events)

Blocked
5 blocks of two

(5 random events)

Experimental Design – latin square designs

	1	2	3	4
1	A	C	B	D
2	C	A	D	B
3	D	B	A	C
4	B	D	C	A

	1	2	3	4
1	A	D	B	C
2	C	A	D	B
3	B	C	A	D
4	D	B	C	A

- A latin square allows the control/balance of external sources of variation - particularly temporal effects or sequence effects in repeated-measures designs (e.g. animals that are learning)
- Latin squares assure that each animal (e.g. columns) see each treatment (e.g. rows) once and each treatment is seen once first, once second…..by an animal
- Subjects and treatments are randomly allocated to columns and rows

Experimental Design

Key elements:

- Control/comparison group
- Replication
- Randomisation

- **Experimenter bias**
  - Much of the experimenter bias has already been eliminated by the above randomisation
  - Blind and double blind experiments are state of the art treatments to reduce the effects even further

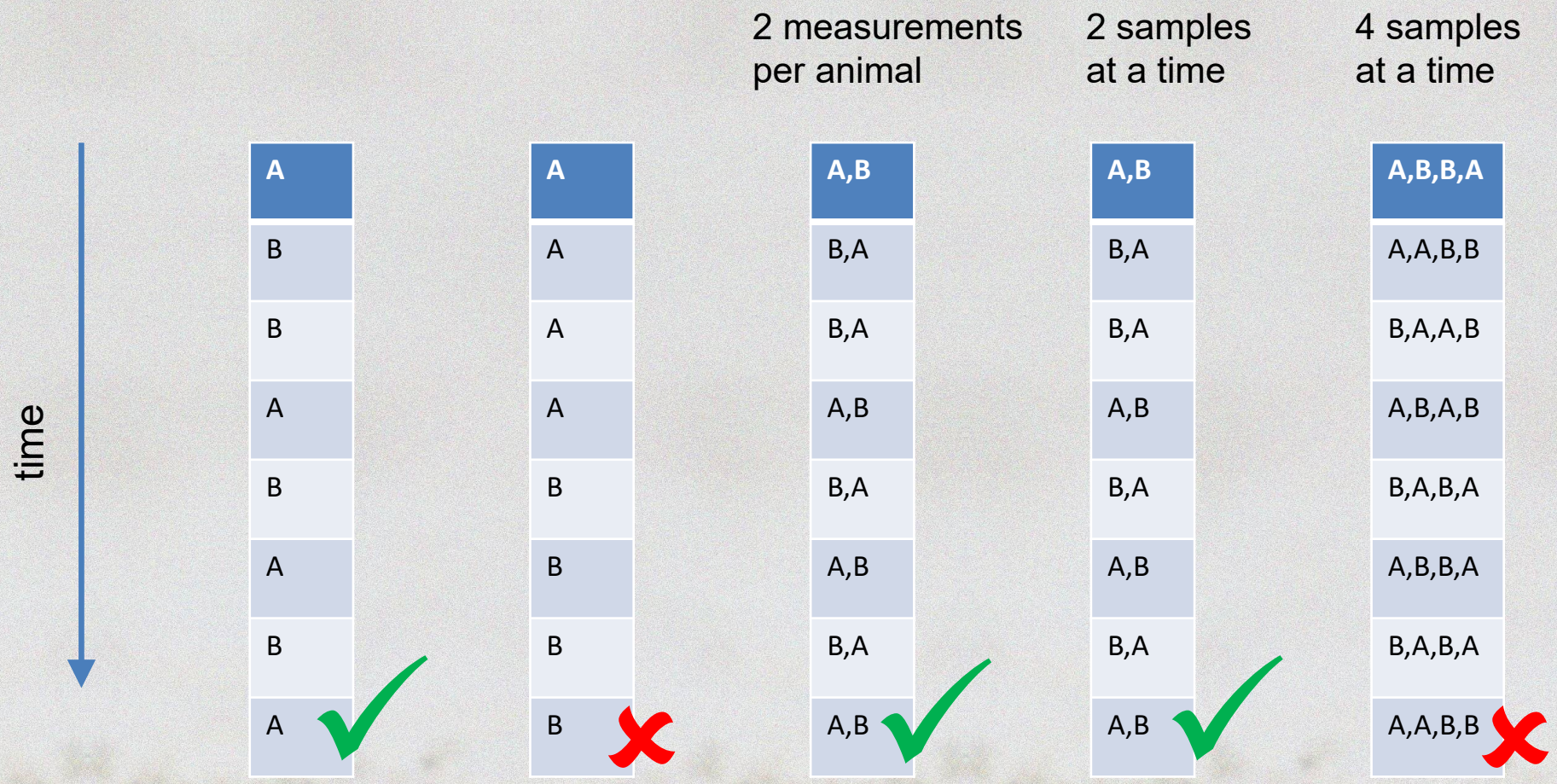
Analysis

The analysis should reflect the experimental design and vice versa.

The most important issue to consider is the independence of replicas

- Blocking
- Repeated measures
- Temporal sequences

All affect the independence of replicas and have to be taken into account to avoid pseudo replication (e.g. repeated measures on the same animal)

Pseudoreplication

Which of these designs is ok and which amounts to pseudoreplication when analysed with a simple t-test comparing A with B? All are sampled over the course of a day (consecutive rows)